Partial Decode–Forward Relaying for the Gaussian Two-Hop Relay Network

Jing Li and Young-Han Kim

***Abstract*—The multicast capacity of the Gaussian two-hop relay network with one source, N relays, and L destinations is studied. It is shown that a careful modification of the partial decode–forward coding scheme, whereby the relays cooperate through degraded sets of message parts, achieves the cutset upper bound within $(1/2)\log N$ bits regardless of the channel gains and power constraints. This scheme improves upon a previous scheme by Chern and Özgür, which is also based on partial decode–forward yet has an unbounded gap from the cutset bound for $L \geq 2$ destinations. When specialized to independent codes among relays, the proposed scheme achieves within $\log N$ bits from the cutset bound. The computation of this relaxation involves evaluating mutual information across $L(N+1)$ cuts out of the total $L2^N$ possible cuts, providing a very simple linear-complexity algorithm to approximate the single-source multicast capacity of the Gaussian two-hop relay network.**

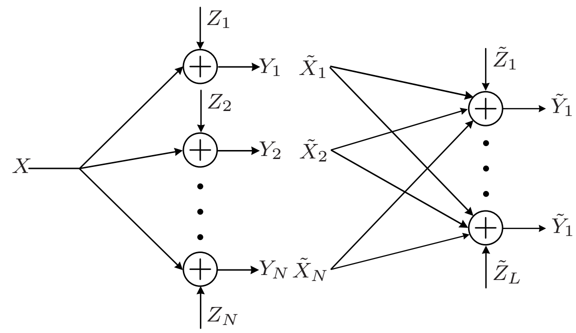

Fig. 1. The Gaussian two-hop relay network.

I. Introduction

Consider the Gaussian two-hop relay network with one source, N relays, and L destinations as depicted in Fig. 1, which can be viewed as a cascade of a broadcast channel (BC) from the source to the relays and multiple multiple access channels (MACs) from the relays to the destinations. The source node wishes to reliably communicate a common message to the L destination nodes with help of the N relays. The special case of $L = 1$, originally introduced by Schein and Gallager [1], [2], is often referred to as the *diamond network*. The capacity is not known in general except for the trivial case of $N = 1$.

The best known capacity upper bound is the cutset bound [3], which is the maximum of the minimum mutual information across all possible cuts that separate the source and the destinations. There are several capacity lower bounds based on different coding schemes. The compress–forward scheme for the 3-node relay channel by Cover and El Gamal [4] has been extended to relay networks in several forms, such as quantize–map–forward (QMF) by Avestimehr, Diggavi, and Tse [5], and noisy network coding (NNC) [6], [7]. The standard analysis [6] shows that when specialized to our two-hop network model in Fig. 1, these coding schemes achieve the cutset bound within $O(N)$ bits for any channel parameters (recall that N is the number of relays).

The material in this paper was presented in part at the 52nd Annual Allerton Conference on Communication, Control, and Computing, Monticello, Illinois, October 2014.

Jing Li is with the School of Telecommunications Engineering, Xidian University, Xi'an, Shaanxi 710071, China (email: jli_8713@stu.xidian.edu.cn) and has been visiting the University of California, San Diego since September 2013. Young-Han Kim is with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093, USA (email: yhk@ucsd.edu).

Recently, Chern and Özgür [8] provided a more refined analysis on the performance of NNC and showed that it achieves within $(1/2)\log 2N(N+1)$ bits from the cutset bound regardless of the number of destinations. In the same paper [8], Chern and Özgür extended the partial decode–forward (PDF) scheme for the relay channel by Cover and El Gamal [4] to the Gaussian diamond network $(L = 1)$. In the PDF scheme by Chern and Özgür, the source broadcasts independent parts of the message to the relays, which in turn recover and forward their corresponding parts to the destination over the MAC. Thus, the Chern–Özgür scheme can achieve the rate characterized by the intersection of the BC capacity region and the MAC capacity region, which can be shown to be within $\log N$ bits from the cutset bound. When there are more than one destination node, however, the gap from the cutset bound becomes unbounded [8, Sec. VI].

In this paper, we develop an alternative extension of partial decode–forward that achieves the cutset bound within $\frac{1}{2}\log N$ bits for any number of destination nodes. In the proposed scheme, the relays decode for multiple message parts based on their respective decoding capabilities (as in the BC with *degraded message sets* [9]) and forward these parts cooperatively (as in the MAC with *degraded message sets* [10], [11]). Thus, the proposed scheme achieves the rate characterized by the intersection of the capacity region of the BC with degraded message sets and the capacity regions of the group of multiple access channels with degraded message sets.

Although this improvement may be viewed at first as an unnatural complication (except for the obvious benefit for achieving higher multicast rates with $L \geq 2$ destinations), it actually yields a simpler characterization of the achievable rate when independent Gaussian random codebooks are used at the relays, which yields a slightly looser but easier-to-compute

$\log N$ approximation of the capacity. A direct computation of the cutset bound as well as of the achievable rates for NNC and the Chern–Özgür PDF scheme requires evaluating mutual information across $L2^N$ different cuts and then taking the minimum, which takes exponential time when directly computed. As an alternative to direct computation, *approximate* computation of the capacity (or the cutset bound) of the single-source single-destination relay network has been proposed by Parvaresh and Etkin [12] based on properties of submodular function minimization, which implies that the capacity of our two-hop network with $L = 1$ can be approximated within $2N$ in polynomial time of $O(LN^6)$ complexity (see also [13]). In this paper, we refine and strengthen the Parvaresh–Etkin approximation result by showing that the achievable rate of our PDF scheme under independent codebooks involves evaluating only $L(N+1)$ cut rates. As a consequence, we develop an explicit algorithm to approximate the capacity as well as the cutset bound within $\log N$ with linear time complexity.

Finally, we evaluate the performance of yet another variant of partial decode–forward for the two-hop relay network. Recently, Lim, Kim, and Kim developed distributed decode–forward, which generalizes partial decode–forward to general noisy networks for multicast [14] and broadcast [15]. As in the case of noisy network coding, a naive analysis of distributed decode–forward results in an achievable rate within $N/2$ bits from the cutset bound. In this paper, we provide a refined analysis that establishes a gap of $(\log N + \frac{1}{2})$ bits from the cutset bound.

The rest of the paper is organized as follows. In the next section, we review basic facts on polymatroids. In Section III, we formally define the capacity of the Gaussian two-hop relay network. In Section IV, we review the cutset upper bound on the capacity, which will be benchmarked throughput. In Section V, we review the Chern–Özgür partial decode–forward scheme for the Gaussian diamond network ($L = 1$). In Section VI, we present our coding scheme for the special case of the diamond network and then extend this result to the general L-destination case. In Section VII, we show the computation of the achievable rate of the relaxed version of our coding scheme involves linear complexity. In Section VIII and put forward the improved analysis of the performance of DDF. Finally, we conclude the paper.

Throughout the paper, we mostly follow the notation in [16]. In particular, we denote $[1:N] := \{1, 2, \cdots, N\}$. The maximum of a finite set is denoted as $\mathcal{J}_{\max} := \max(\mathcal{J})$. A tuple of random variables is denoted as $X(\mathcal{J}) := (X_j : j \in \mathcal{J})$. The Gaussian capacity function is defined as $\mathsf{C}(x) := (1/2)\log(1+x)$.

II. Mathematical Preliminaries

Let $\phi : 2^{[1:N]} \to [0, \infty)$ be a set function satisfying

1) $\phi(\emptyset) = 0$,
2) $\phi(\mathcal{J}) \le \phi(\mathcal{K})$ if $\mathcal{J} \subseteq \mathcal{K}$, and
3) $\phi(\mathcal{J} \cap \mathcal{K}) + \phi(\mathcal{J} \cup \mathcal{K}) \le \phi(\mathcal{J}) + \phi(\mathcal{K})$.

Then the polyhedron

$$\mathscr{P}(\phi) := \Big\{(x_1, \cdots, x_N) \in [0,\infty)^N : \sum_{j\in\mathcal{J}} x_j \le \phi(\mathcal{J}),\ \mathcal{J} \subseteq [1:N]\Big\}$$

is said to be a *polymatroid* (associated with ϕ); see, for example, [17].

Example 1. For any random tuple $(X_1, \ldots, X_N, Y)$ such that $X_1, \ldots, X_N$ are mutually independent, the set of rate tuples $(R_1, \ldots, R_N)$ satisfying

$$\sum_{j\in\mathcal{J}} R_j \le I(X(\mathcal{J}); Y \,|\, X(\mathcal{J}^c))$$

is a polymatroid [11, Lemma 3.1]. In particular, if $X_j \sim \mathrm{N}(0, S_j)$, $j \in [1:N]$, and $Y = \sum_{j=1}^N X_j + Z$, where $X_1, \ldots, X_N$ and $Z \sim \mathrm{N}(0,1)$ are mutually independent, then the set of rate tuples $(R_1, \ldots, R_N)$ satisfying

$$\sum_{j\in\mathcal{J}} R_j \le \mathsf{C}\Bigg(\sum_{j\in\mathcal{J}} S_j\Bigg)$$

is a polymatroid.

Example 2. Let $\Phi : [1:N] \to [0,\infty)$ be nondecreasing and define $\phi : 2^{[1:N]} \to [0, \infty)$ by

$$\phi(\mathcal{J}) = \begin{cases} 0, & \mathcal{J} = \emptyset, \\ \Phi(\mathcal{J}_{\max}), & \text{otherwise.} \end{cases}$$

Then it can be readily shown that $\mathscr{P}(\phi)$ is a polymatroid characterized by active inequalities

$$\sum_{j=1}^k x_j \le \phi([1:k]) = \Phi(k), \quad k \in [1:N].$$

In particular, for any random tuple $(X_1, \ldots, X_N, Y)$, the set of rate tuples $(R_1, \ldots, R_N)$ satisfying

$$\sum_{j=1}^k R_j \le I(X^k; Y \,|\, X_{k+1}^N)$$

is a polymatroid.

The following well-known result is pivotal in our discussion.

Lemma 1 (Edmonds's polymatroid intersection theorem [18])**.** *If $\mathscr{P}(\phi)$ and $\mathscr{P}(\psi)$ are two polymatroids, then*

$$\max\Big\{\sum_{j=1}^N x_j : (x_1, \cdots, x_N) \in \mathscr{P}(\phi) \cup \mathscr{P}(\psi)\Big\} = \min_{\mathcal{J}\subseteq[1:N]} \big[\phi(\mathcal{J}) + \psi(\mathcal{J}^c)\big].$$

III. Formal Definition of Capacity

Recall the Gaussian two-hop relay network model depicted in Fig. 1. The received signals at the relays corresponding to the signal X from the source node are

$$Y_j = g_j X + Z_j, \quad j \in [1:N],$$

where $g_1, \dots, g_N$ are the channel gains from the source to relay nodes 1 through N, respectively, and $Z_1, \dots, Z_N$ are independent $\mathrm{N}(0,1)$ noise components. We assume without loss of generality that

$$|g_1| \ge |g_2| \ge \cdots \ge |g_N|. \tag{1}$$

Similarly, the received signals at the destinations corresponding to the signals $\tilde{X}_1, \dots, \tilde{X}_N$ transmitted from the relays are

$$\tilde{Y}_d = \sum_{j=1}^{N} \tilde{g}_{dj} \tilde{X}_j + \tilde{Z}_d, \quad d \in [1:L],$$

where $\tilde{g}_{dj}$, $j \in [1:N]$, $d \in [1:L]$, denote the channel gain from relay node j to destination node d, and $\tilde{Z}_1, \dots, \tilde{Z}_L$ are independent $\mathrm{N}(0,1)$ noise components. The first (source-to-relays) hop of the network can be viewed as a Gaussian broadcast channel, while the second (relays-to-destinations) hop of the network can be viewed as multiple Gaussian multiple access channels. All nodes are subject to (expected) average power constraint P, and we denote by $S_j = g_j^2 P$ and $\tilde{S}_{dj} = \tilde{g}_{dj}^2 P$ the received signal-to-noise ratios (SNRs) at the relays and the receivers, respectively.

We define a $(2^{nR}, n)$ code for a Gaussian two-hop relay network by

- a message set $[1:2^{nR}]$,
- an source encoder that assigns a codeword $x^N(m)$ to each message $m \in [1:2^{nR}]$,
- a set of relay encoders, where encoder $j \in [1:N]$ assigns a symbol $\tilde{x}_{ji}(y_j^{i-1})$ to each past received sequence y_j^{i-1} for each transmission time $i \in [1:n]$, and
- a set of decoders, where decoder $d \in [1:L]$ assigns an estimate $\hat{m}_d$ or an error message e to each received sequence $\tilde{y}_d^N$.

We assume that the message M is uniformly distributed over the message set. The average probability of error is defined as $P_e^{(n)} = \mathsf{P}\{\hat{M}_d \neq M \text{ for some } d \in [1:L]\}$. A rate R is said to be achievable for the Gaussian two-hop relay network if there exists a sequence of $(2^{nR}, n)$ codes such that $\lim_{n\to\infty} P_e^{(n)} = 0$. The capacity C is defined as the supremum of all achievable rates.

When $N = 1$, the capacity is

$$C = \min\Big\{\mathsf{C}(S_1),\, \min_d \mathsf{C}(\tilde{S}_{d1})\Big\}.$$

For $N \ge 2$, however, no computable characterization of the capacity is known even when $L = 1$. In subsequent sections, we present bounds on the capacity and establish their closeness.

IV. The Cutset Bound on the Capacity

Since the network consists of two noninteracting channel layers, the cutset bound [3] on the capacity of a general noisy network can be simplified as

$$\begin{aligned} C &\le R_{\mathrm{CS}} \\ &:= \sup_F \min_{d,\mathcal{J}} \big[I(X; Y(\mathcal{J}^c)) + I(\tilde{X}(\mathcal{J}); \tilde{Y}_d | \tilde{X}(\mathcal{J}^c)) \big], \end{aligned} \tag{2}$$

where the supremum is over all joint distributions $F(x)F(\tilde{x}^N)$ satisfying $\mathsf{E}(X^2) \le P$ and $\mathsf{E}(\tilde{X}_j^2) \le P$, $j \in [1:N]$, the minimum is over all $d \in [1:L]$ and $\mathcal{J} \subseteq [1:N]$, and $\mathcal{J}^c$ denotes $[1:N] \setminus \mathcal{J}$. By the maximum differential entropy lemma (see, for example, [16, Section 2.2]), the supremum in (2) is attained by Gaussian X and jointly Gaussian $(\tilde{X}_1, \dots, \tilde{X}_N)$. By switching the order of the supremum (over Gaussian distributions) and the minimum, the cutset bound is further upper bounded as

$$\begin{aligned} R_{\mathrm{CS}} &\le \sup_{F(\tilde{x}^N)} \min_{d,\mathcal{J}} \sup_{F(x)} \big[I(X; Y(\mathcal{J}^c)) + I(\tilde{X}(\mathcal{J}); \tilde{Y}_d | \tilde{X}(\mathcal{J}^c)) \big] \\ &= \sup_{F(\tilde{x}^N)} \min_{d,\mathcal{J}} \Bigg[\mathsf{C}\Bigg(\sum_{j \in \mathcal{J}^c} S_j \Bigg) + I(\tilde{X}(\mathcal{J}); \tilde{Y}_d | \tilde{X}(\mathcal{J}^c)) \Bigg] \end{aligned} \tag{3}$$

$$\begin{aligned} &\le \min_{d,\mathcal{J}} \sup_{F(\tilde{x}^N)} \Bigg[\mathsf{C}\Bigg(\sum_{j \in \mathcal{J}^c} S_j \Bigg) + I(\tilde{X}(\mathcal{J}); \tilde{Y}_d | \tilde{X}(\mathcal{J}^c)) \Bigg] \\ &\le \min_{d,\mathcal{J}} \Bigg[\mathsf{C}\Bigg(\sum_{j \in \mathcal{J}^c} S_j \Bigg) + \mathsf{C}\Bigg(\Bigg(\sum_{j \in \mathcal{J}} \sqrt{\tilde{S}_{dj}} \Bigg)^2 \Bigg) \Bigg]. \end{aligned} \tag{4}$$

Note that direct computation of the cutset bound in (3) for a fixed distribution or its relaxation in (4) involves evaluation of the minimum rate over the combination of 2^N choices of $\mathcal{J}$ and L choices of d, that is, the total $L2^N$ cuts that separate the source and the destinations.

V. The Chern–Özgür Partial Decode–Forward Scheme for the Gaussian Diamond Network

In the partial decode–forward scheme by Chern and Özgür [8] (see also [19]), which was developed mainly for the case $N = 1$, the source node divides the message M into N independent parts $M_1, \dots, M_N$ (rate splitting), relay j recovers M_j and forwards it (decode–forward), and the destination node forms the estimates of $M_1, \dots, M_N$ and thus of M itself; see Fig. 2. This scheme is implemented over two hops in a block Markov fashion, and the achievable rate can be characterized as

$$R_{\mathrm{PDF}} = \max\Bigg\{ \sum_{j=1}^{N} R_j \colon (R_1, \dots, R_N) \in \mathscr{R}_{\mathrm{BC}} \cap \mathscr{R}_{\mathrm{MAC}} \Bigg\}. \tag{5}$$

Here $\mathscr{R}_{\mathrm{BC}}$ is the capacity region of the standard N-receiver Gaussian broadcast channel with SNRs $S_1, \dots, S_N$, that is,

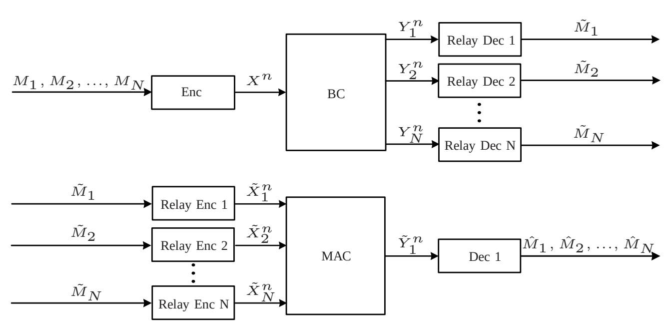

Fig. 2. The Chern–Özgür partial decode–forward coding scheme for $L = 1$.

the set of rate tuples $(R_1, \dots, R_N)$ such that

$$R_j \le \mathsf{C}\left(\frac{\alpha_j S_j}{\sum_{k=1}^{j-1} \alpha_k S_k + 1}\right), \quad j \in [1:N], \tag{6}$$

for some $(\alpha_1, \dots, \alpha_N)$ satisfying $\alpha_j \ge 0$, $j \in [1:N]$, and $\sum_{j=1}^N \alpha_j = 1$, which, by the BC–MAC duality [20], can be written as the set of rate pairs $(R_1, \dots, R_N)$ such that

$$\sum_{j\in\mathcal{J}} R_j \le \mathsf{C}\left(\sum_{j\in\mathcal{J}} \beta_j S_j\right), \quad \mathcal{J} \subseteq [1:N], \tag{7}$$

for some $(\beta_1, \dots, \beta_N)$ satisfying $\beta_j \ge 0$, $j \in [1:N]$, and $\sum_{j=1}^N \beta_j = 1$. In (5), $\mathscr{R}_{\mathrm{MAC}}$ is the capacity region of the standard N-sender Gaussian multiple access channel with SNRs $\tilde{S}_{11}, \dots, \tilde{S}_{1N}$, i.e., the set of rate tuples $(R_1, \dots, R_N)$ such that

$$\sum_{j\in\mathcal{J}} R_j \le \mathsf{C}\left(\sum_{j\in\mathcal{J}} \tilde{S}_{1j}\right), \quad \mathcal{J} \subseteq [1:N].$$

Note that the region $\mathscr{R}_{\mathrm{MAC}}$ is a polymatroid (cf. Example 1), but the region $\mathscr{R}_{\mathrm{BC}}$ is not in general. Consequently, the maximum sum-rate of the intersection of the two regions, characterized by (5), is rather cumbersome to calculate. Chern and Özgür set $\beta_j \equiv 1/N$ in (7) to obtain a *polymatroidal* inner bound on $\mathscr{R}_{\mathrm{BC}}$ characterized by

$$\sum_{j\in\mathcal{J}} R_j \le \mathsf{C}\left(\frac{1}{N}\sum_{j\in\mathcal{J}} S_j\right), \quad \mathcal{J} \subseteq [1:N]. \tag{8}$$

Now by (5) and Edmonds's polymatroid intersection theorem with

$$\phi(\mathcal{J}) = \mathsf{C}\left(\sum_{j\in\mathcal{J}} \tilde{S}_{1j}\right),$$
$$\psi(\mathcal{J}) = \mathsf{C}\left(\frac{1}{N}\sum_{j\in\mathcal{J}} S_j\right),$$

the corresponding (lower bound on the) achievable rate is

$$\begin{aligned} R_{\mathrm{PDF}} &\ge \min_{\mathcal{J}\subseteq[1:N]} \left[\phi(\mathcal{J}) + \psi(\mathcal{J}^c)\right] \\ &= \min_{\mathcal{J}\subseteq[1:N]} \left[\mathsf{C}\left(\frac{1}{N}\sum_{j\in\mathcal{J}^c} S_j\right) + \mathsf{C}\left(\sum_{j\in\mathcal{J}} \tilde{S}_{1j}\right)\right]. \end{aligned} \tag{9}$$

By comparing this rate with the capacity upper bound in (4), we observe that the gaps for the two terms, both due to the lack of coherent cooperation, are bounded uniformly as

$$\mathsf{C}\left(\sum_{j\in\mathcal{J}^c} S_j\right) - \mathsf{C}\left(\frac{1}{N}\sum_{j\in\mathcal{J}^c} S_j\right) \le \frac{1}{2}\log N, \tag{10}$$

$$\mathsf{C}\left(\left(\sum_{j\in\mathcal{J}} \sqrt{\tilde{S}_{1j}}\right)^2\right) - \mathsf{C}\left(\sum_{j\in\mathcal{J}} \tilde{S}_{1j}\right) \le \frac{1}{2}\log N. \tag{11}$$

In conclusion, the gap between the achievable rate of the Chern–Özgür partial decode–forward scheme and the cutset bound is upper bounded as

$$\Delta_{\mathrm{PDF}} := R_{\mathrm{CS}} - R_{\mathrm{PDF}} \le \log N,$$

regardless of S_j and $\tilde{S}_{1k}$, $j, k \in [1:N]$.

VI. The Proposed Partial Decode–Forward Scheme

We propose a modified version of the Chern–Özgür partial decode–forward scheme as depicted in Fig. 3. Here, the relays recover degraded sets of the message parts in the natural order—recall the assumption on the channel gains in (1)—say, relay 1 recovers $(M_1, \dots, M_N)$, relay 2 recovers $(M_2, \dots, M_N)$, relay 3 recovers $(M_3, \dots, M_N)$, and so on. The relays then cooperatively communicate these message parts to each destination as in the multiple access channel with degraded message sets [10], [11].

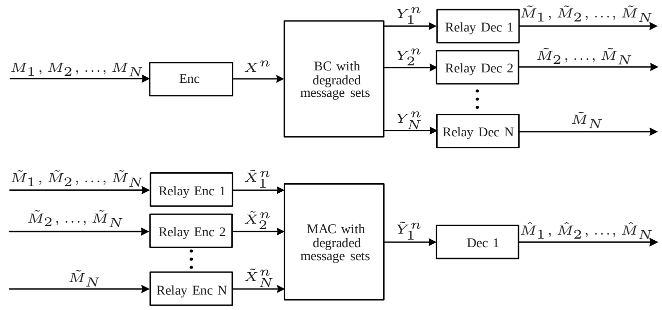

Fig. 3. The proposed partial decode–forward scheme for $L = 1$.

A. The Diamond Network

For simplicity of exposition, we first consider the case $L = 1$. The achievable rate of the proposed scheme can be characterized as

$$R'_{\mathrm{PDF}} = \max\left\{\sum_{j=1}^N R_j \colon (R_1, \dots, R_N) \in \mathscr{R}'_{\mathrm{BC}} \cap \mathscr{R}'_{\mathrm{MAC}}\right\},$$

where $\mathscr{R}'_{\mathrm{BC}}$ is the capacity region of the standard N-receiver Gaussian broadcast channel (BC) with degraded message sets and $\mathscr{R}'_{\mathrm{MAC}}$ is the capacity region of the N-sender Gaussian multiple access channel (MAC) with degraded message sets. Since the broadcast channel is degraded in the order of $1 \to 2 \to \cdots \to N$, $\mathscr{R}'_{\mathrm{BC}} = \mathscr{R}_{\mathrm{BC}}$ as in (6). The capacity region of the multiple access channel with degraded message sets [10], [11] consists of all rate tuples $(R_1, \dots, R_N)$ such that

$$\sum_{j=1}^k R_j \le I(\tilde{X}^k; \tilde{Y}_1 | \tilde{X}_{k+1}^N), \quad k \in [1:N], \tag{12}$$

for some $F(\tilde{x}^N)$ such that $\mathsf{E}(\tilde{X}_j^2) \le P$, $j \in [1:N]$. Again by the maximum differential entropy lemma, there is no loss of generality in setting $(\tilde{X}_1, \dots, \tilde{X}_N)$ to be jointly Gaussian in (12).

In order to obtain a lower bound on R'_{PDF}, we follow the same approach [8], [19] as reviewed in the previous section and use the polymatroidal inner bound on $\mathscr{R}'_{\mathrm{BC}}$ in (8). As for $\mathscr{R}'_{\mathrm{MAC}}$, we note that the region in (12) is a polymatroid for a fixed $F(\tilde{x}^N)$; cf. Example 2. Thus, by Edmonds's polymatroid intersection theorem with

$$\phi(\mathcal{J}) = I(\tilde{X}^{\mathcal{J}_{\max}}; \tilde{Y}_1 | \tilde{X}_{\mathcal{J}_{\max}+1}^N), \tag{13}$$
$$\psi(\mathcal{J}) = \mathsf{C}\left(\frac{1}{N}\sum_{j\in\mathcal{J}} S_j\right),$$

the achievable rate of the proposed scheme is lower bounded as

$$R'_{\mathrm{PDF}} \geq \sup_{F} \min_{\mathcal{J}\subseteq[1:N]} \big[\psi(\mathcal{J}^c) + \phi(\mathcal{J})\big], \tag{14}$$

where the supremum is over all jointly Gaussian $\tilde{X}^N$ satisfying $\mathsf{E}(\tilde{X}_j^2) \leq P$, $j \in [1:N]$. Since for each $\mathcal{J} \subseteq [1:N]$ with $\mathcal{J}_{\max} = k$,

$$\begin{aligned}\psi(\mathcal{J}^c) + \phi(\mathcal{J}) &\geq \psi([1:k]^c) + \phi(\mathcal{J}) \\ &= \psi([k+1:N]) + \phi([1:k]),\end{aligned}$$

the minimum in (14) is attained by $\mathcal{J} = \emptyset$ or $\mathcal{J} = [1:k]$ for some k. Thus,

$$R'_{\mathrm{PDF}} \geq \sup_{F} \min_{k\in[0:N]} \bigg[\mathsf{C}\bigg(\frac{1}{N}\sum_{j=k+1}^{N} S_j\bigg) + I(\tilde{X}^k; \tilde{Y}_1 | \tilde{X}_{k+1}^N)\bigg]. \tag{15}$$

In comparison, by restricting $\mathcal{J}$ to be of the form $[1:k]$ in (3), the cutset upper bound can be relaxed as

$$R_{\mathrm{CS}} \leq \sup_{F} \min_{k\in[0:N]} \bigg[\mathsf{C}\bigg(\sum_{j=k+1}^{N} S_j\bigg) + I(\tilde{X}^k; \tilde{Y}_1 | \tilde{X}_{k+1}^N)\bigg]. \tag{16}$$

By comparing (15) and (16), we establish the following.

Proposition 1. *The gap between the achievable rate of the proposed partial decode–forward scheme and the cutset bound is upper bounded as*

$$\Delta'_{\mathrm{PDF}} := R_{\mathrm{CS}} - R'_{\mathrm{PDF}} \leq \frac{1}{2}\log N,$$

regardless of S_j and $\tilde{S}_{1k}$, $j,k \in [1:N]$.

B. The General Two-Hop Network

The advantage of the modified partial decode–forward coding scheme is fully realized when there are multiple destinations ($L \geq 2$), in which case the Chern–Özgür scheme has an unbounded gap from the capacity [8, Sec. VI]. Recall from Fig. 3 that in the proposed partial decode–forward scheme, the message parts are communication over a cascade of a BC (with degraded message sets) and multiple MACs with degraded message sets. The achievable rate can be thus characterized as

$$R'_{\mathrm{PDF}} = \max\bigg\{\sum_{j=1}^{N} R_j \colon (R_1,\ldots,R_N) \in \mathscr{R}'_{\mathrm{BC}} \cap \mathscr{R}'_{\mathrm{MMAC}}\bigg\},$$

where $\mathscr{R}'_{\mathrm{MMAC}}$ is the set of rate tuples $(R_1,\ldots,R_N)$ such that

$$\sum_{j=1}^{k} R_j \leq \min_{d\in[1:L]} I(\tilde{X}^k; \tilde{Y}_d | \tilde{X}_{k+1}^N), \quad k \in [1:N], \tag{17}$$

for some jointly Gaussian $\tilde{X}^N$ with $\mathsf{E}(\tilde{X}_j^2) \leq P$, $j \in [1:N]$, which is identical to the capacity region of the N-sender L-state Gaussian compound MAC with degraded message sets.

We can now proceed in the exactly same manner as in the single-destination case, except that in place of (13) we have another polymatroid

$$\phi(\mathcal{J}) = \min_{d\in[1:L]} I(\tilde{X}^{\mathcal{J}_{\max}}; \tilde{Y}_d | \tilde{X}_{\mathcal{J}_{\max}+1}^N).$$

Consequently, we can lower bound the achievable rate of the scheme as

$$\begin{aligned}&R'_{\mathrm{PDF}} \\ &\geq \sup_{F} \min_{d\in[1:L]} \min_{k\in[0:N]} \bigg[\mathsf{C}\bigg(\frac{1}{N}\sum_{j=k+1}^{N} S_j\bigg) + I(\tilde{X}^k; \tilde{Y}_d | \tilde{X}_{k+1}^N)\bigg].\end{aligned} \tag{18}$$

In comparison,

$$\begin{aligned}&R_{\mathrm{CS}} \\ &\leq \sup_{F} \min_{d\in[1:L]} \min_{k\in[0:N]} \bigg[\mathsf{C}\bigg(\sum_{j=k+1}^{N} S_j\bigg) + I(\tilde{X}^k; \tilde{Y}_d | \tilde{X}_{k+1}^N)\bigg].\end{aligned}$$

This establishes the following.

Theorem 1. *The gap between the achievable rate of the proposed partial decode–forward scheme and the cutset bound is upper bounded as*

$$\Delta'_{\mathrm{PDF}} = R_{\mathrm{CS}} - R'_{\mathrm{PDF}} \leq \frac{1}{2}\log N,$$

regardless of the SNRs S_j and $\tilde{S}_{dk}$, $j,k \in [1:N]$, $d \in [1:L]$, and the number of destinations $L = 1, 2, \ldots$.

A few remarks are in order.

1) When $\mathscr{R}'_{\mathrm{MMAC}} \subseteq \mathscr{R}'_{\mathrm{BC}}$ (which is the case, for example, if $|g_N| \geq \min_d \sum_{j=1}^N |\tilde{g}_{dj}|$), the proposed coding scheme actually achieves the capacity

$$C = \min_{d\in[1:L]} \mathsf{C}\bigg(\bigg(\sum_{j=1}^{N}\sqrt{\tilde{S}_{dj}}\bigg)^2\bigg).$$

In this case, the coding scheme simplifies to a simple decode–forward scheme, whereby every relay recovers the message M and coherently forwards it.

2) At the other extreme, when $\mathscr{R}'_{\mathrm{BC}} \subseteq \mathscr{R}'_{\mathrm{MMAC}}$ (which is the case, for example, if $|g_1| \leq \min_d |\tilde{g}_{d1}|$), the maximum achievable rate of the proposed coding scheme is

$$R'_{\mathrm{PDF}} = \mathsf{C}(S_1).$$

Note that this rate is achieved trivially by using only the best relay (relay 1) and keeping the other relays idle, yet the gap from the capacity is no more than $(1/2)\log N$.

The performance difference between the Chern–Özgür PDF scheme and the proposed PDF scheme is best illustrated by the following example taken from [8, Sec. VI].

Example 3. Consider the Gaussian two-hop relay network with 2 relays and 2 destinations as depicted in Fig. 4, where the coefficients indicate the corresponding channel gains. The cutset bound is bounded as

$$\mathsf{C}(a^2P) \leq R_{\mathrm{CS}} \leq \mathsf{C}((a+\sqrt{a})^2P),$$

where the lower bound follows by setting $X, \tilde{X}_1, \tilde{X}_2$ to be independent $\mathrm{N}(0,P)$ in (2) and the upper bound follows by considering only the broadcast cut. The achievable rate of the PDF scheme by Chern and Özgür is

$$R_{\mathrm{PDF}} = \mathsf{C}(aP),$$

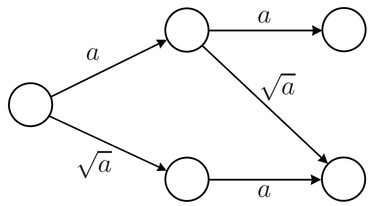

Fig. 4. An example network.

which has an arbitrarily large gap from the cutset bound as $a \to \infty$. In comparison, the achievable rate of the proposed PDF scheme is lower bounded as

$$R'_{\mathrm{PDF}} \geq \mathsf{C}\left(\frac{(a+a^2)P}{2}\right),$$

which is within 1 bit from the cutset bound.

VII. Linear-Complexity Capacity Approximation

Computation of the achievable rate in (18) requires maximization over all Gaussian input distributions F. We now restrict the distribution to be independent and identically distributed $\tilde{X}_j \sim \mathrm{N}(0, P)$, $j \in [1:N]$. This can be interpreted as a more practical coding scheme in which the relays use independent Gaussian codebooks and transmit codewords non-coherently. The achievable rate of the scheme is lower bounded by

$$R''_{\mathrm{PDF}} \geq \min_{d\in[1:L]} \min_{k\in[0:N]} \left[\mathsf{C}\left(\frac{1}{N}\sum_{j=k+1}^{N} S_j\right) + \mathsf{C}\left(\sum_{j=1}^{k} \tilde{S}_{dj}\right)\right]. \tag{19}$$

In comparison, starting with (4) and following the same argument as before, we can relax the cutset upper bound as

$$R_{\mathrm{CS}} \leq \min_{d\in[1:L]} \min_{k\in[0:N]} \left[\mathsf{C}\left(\sum_{j=k+1}^{N} S_j\right) + \mathsf{C}\left(\left(\sum_{j=1}^{k} \sqrt{\tilde{S}_{dj}}\right)^2\right)\right]. \tag{20}$$

Thus, by (10) and (11), the capacity is approximated uniformly by $\log N$. Moreover, the computation of (19) or (20) involves computing Gaussian capacity functions for $L(N+1)$ cuts, which is a significant savings from the directed computation of the cutset bound with all $L2^N$ possible cuts as in (2).

We summarize this result as follows.

Proposition 2. *The capacity of the Gaussian two-hop network is bounded as*

$$C \geq \min_{d\in[1:L]} \min_{k\in[0:N]} \left[\mathsf{C}\left(\frac{1}{N}\sum_{j=k+1}^{N} S_j\right) + \mathsf{C}\left(\sum_{j=1}^{k} \tilde{S}_{dj}\right)\right],$$

$$C \leq \min_{d\in[1:L]} \min_{k\in[0:N]} \left[\mathsf{C}\left(\sum_{j=k+1}^{N} S_j\right) + \mathsf{C}\left(\left(\sum_{j=1}^{k} \sqrt{\tilde{S}_{dj}}\right)^2\right)\right],$$

where the gap between the lower and upper bounds is no greater than $\log N$ *for any* S_j *and* $\tilde{S}_{dk}$, $j, k \in [1:N]$, $d \in [1:L]$, *and any* L. *Moreover, both bounds can be computed in* $O(LN)$ *complexity.*

These bounds yields a simple approximate expression for the capacity.

Proposition 3.

$$C = \min_{d\in[1:L]} \min_{k\in[0:N]} \left[\mathsf{C}\left(\sum_{j=k+1}^{N} S_j\right) + \mathsf{C}\left(\sum_{j=1}^{k} \tilde{S}_{dj}\right)\right] \pm \frac{1}{2}\log N.$$

VIII. Distributed Decode–Forward

In this section, we consider the distributed decode–forward (DDF) coding scheme in [14], which is an extension of partial decode–forward to general multicast networks. In particular, the rate achieved by DDF for our two-hop network is characterized [14] as

$$\begin{aligned} R_{\mathrm{DDF}} = \sup_F \min_{d,\mathcal{J}} \Big[& I(X, \tilde{X}(\mathcal{J}); U(\mathcal{J}^c), \tilde{Y}_d \,|\, \tilde{X}(\mathcal{J}^c)) \\ & - \sum_{k\in\mathcal{J}^c} I(U_k; X, \tilde{X}^N | Y_k)\Big], \end{aligned} \tag{21}$$

where the supremum is over all distributions of the form $(\prod_{k=1}^{N} F(\tilde{x}_k))F(x|\tilde{x}^N)F(u^N|x,\tilde{x}^N)$ satisfying $\mathsf{E}(X^2) \leq P$ and $\mathsf{E}(\tilde{X}_j^2) \leq P$, $j \in [1:N]$. By setting X and $\tilde{X}_j$ to be i.i.d. $\mathrm{N}(0, P)$ and

$$U_j = Y_j - Z_j + \hat{Z}_j, \quad j \in [1:N], \tag{22}$$

where $\hat{Z}_j \sim \mathrm{N}(0, 1)$, $j \in [1:N]$, are independent of each other and of $(\tilde{X}^N, Y^N)$, it can be shown [14] that the gap between the achievable rate in (21) and the cutset bound in (4) is no greater than $N/2$.

We now exploit the layered structure of the network to improve this $O(N)$ gap to $O(\log N)$. Following a similar (and in some sense dual) development for noisy network coding in [8], we set $\hat{Z}_j \sim \mathrm{N}(0, N)$ in (22). Then, the first term of (21) becomes

$$\begin{aligned} & I(X, \tilde{X}(\mathcal{J}); U(\mathcal{J}^c), \tilde{Y}_d | \tilde{X}(\mathcal{J}^c)) \\ & \quad \overset{(a)}{=} I(X; U(\mathcal{J}^c)) + I(\tilde{X}(\mathcal{J}); \tilde{Y}_d | \tilde{X}(\mathcal{J}^c)) \\ & \quad = \mathsf{C}\left(\frac{1}{N}\sum_{j\in\mathcal{J}^c} S_j\right) + \mathsf{C}\left(\sum_{j\in\mathcal{J}} \tilde{S}_{dj}\right), \end{aligned}$$

where (a) follows by the independence of (X, U^N) and $\tilde{X}^N$ and the layered structure of the network. For $k \in [1:N]$, each summand in the second term of (21) becomes

$$\begin{aligned} I(U_k; X, \tilde{X}^N | Y_k) &= \frac{1}{2}\log\left(\frac{1 + \left(1 + \frac{1}{N}\right)S_k}{1 + S_k}\right) \\ &\leq \frac{1}{2}\log\left(1 + \frac{1}{N}\right) \\ &\leq \frac{1}{2N}. \end{aligned}$$

Hence,

$$R_{\mathrm{DDF}} \geq \min_{d,\mathcal{J}} \left[\mathsf{C}\left(\frac{1}{N}\sum_{j\in\mathcal{J}^c} S_j\right) + \mathsf{C}\left(\sum_{j\in\mathcal{J}} \tilde{S}_{dj}\right) - \frac{|\mathcal{J}^c|}{2N}\right]. \tag{23}$$

Comparing this achievable rate in (23) and the cutset bound in (4) establishes the following.

Proposition 4. *The gap between the achievable rate of the distributed decode–forward scheme and the cutset bound is upper bounded as*

$$\Delta_{\text{DDF}} = R_{\text{CS}} - R_{\text{DDF}} \leq \log N + \frac{1}{2},$$

regardless of the SNRs S_j and $\tilde{S}_{dk}$, $j, k \in [1:N]$, $d \in [1:L]$, and the number of destinations $L = 1, 2, \ldots$.

IX. Concluding Remarks

Multiple coding schemes achieve the multicast capacity of the two-hop Gaussian network with one source, N relays, and L destinations within $O(\log N)$, including:

1) Noisy network coding (see [19, Th. 3.1])
2) Distributed decode–forward (Prop. 4 in the current paper)
3) Partial decode–forward (see [19, Th. 3.3] for $L = 1$)
4) Partial decode–forward with degraded message sets (Th. 1 in the current paper).

Among these, the fourth scheme, which is the main contribution of the paper, achieves the tightest gap of $(1/2)\log N$ from the cutset bound. Moreover, a simple lower bound on its achievable rate can be expressed as the minimum of $L(N+1)$ cut rates, providing a sharp approximation of the capacity that can be computed in $O(LN)$ complexity. While it remains to be seen whether this linear-complexity approximation can be alternatively established via algebraic or combinatorial techniques, it is refreshing to note that the best *computational* result is obtained by a purely information-theoretic argument, based directly on a simple coding scheme.

References

[1] B. Schein and R. G. Gallager, "The Gaussian parallel relay channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Sorrento, Italy, Jun. 2000, p. 22.

[2] B. Schein, "Distributed coordonation in network information theory," Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, USA, Oct. 2001.

[3] A. El Gamal, "On information flow in relay networks," in *Proc. IEEE National Telecomm. Conf.*, New Orleans, LA, Nov. 1981, vol. 2, pp. D4.1.1–D4.1.4.

[4] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[5] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.

[6] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3132–3152, May 2011.

[7] M. H. Yassaee and M. R. Aref, "Slepian–Wolf coding over cooperative relay networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3462–3482, June 2011.

[8] B. Chern and A. Özgür, "Achieving the capacity of the n-relay Gaussian diamond network within $\log n$ bits," in *Proc. IEEE Inf. Theory Workshop*, Lausanne, Switzerland, Sep 2012, pp. 377–380.

[9] P. P. Bergmans, "A simple converse for broadcast channels with additive white Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 279–280, 1974.

[10] V. V. Prelov, "Transmission over a multiple-access channel with a special source hierarchy," *Probl. Peredachi Inf.*, vol. 20, no. 4, pp. 3–10, 1984.

[11] T. S. Han, "The capacity region of general multiple-access channel with certain correlated sources," *Inf. Control*, vol. 40, no. 1, pp. 37–60, 1979.

[12] F. Parvaresh and R. H. Etkin, "On computing the capacity of relay networks in polynomial time," in *Proc. IEEE Int. Symp. Inf. Theory*, Palo Alto, Carlifornia, July/Aug. 2011, pp. 1342–1346.

[13] S. Brahma, A. Sengupta, and C. Fragouli, "Efficient subnetwork selection in relay networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, June/July 2014, pp. 1927–1931.

[14] S. Lim, K. T. Kim, and Y.-H. Kim, "Distribued decode–forward for multicasting," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, June/July 2014, pp. 636–640.

[15] ——, "Distribued decode–forward for broadcast," in *Proc. IEEE Inf. Theory Workshop*, Hobart, TAS, November 2014, pp. 556–560.

[16] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge: Cambridge University Press, 2011.

[17] A. Schrijver, *Combinatorial Optimization.* 3 vols. Berlin: Springer-Verlag, 2003.

[18] J. Edmonds, "Submodular functions, matroids, and certain polyhedra," in *Combinatorial Structures and Their Applications*, Gordon and Breach, New York, 1970, pp. 69–87.

[19] B. Chern and A. Özgür, "Achieving the capacity of the n-relay Gaussian diamond network within $\log n$ bits," *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7708–7718, Dec. 2014.

[20] S. Vishwanath, N. Jindal, and A. J. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, 2003.